Supporting Online Material for

A Systems Approach to Measuring the Binding Energy Landscapes of Transcription Factors

Sebastian J. Maerkl and Stephen R. Quake*

*To whom correspondence should be addressed. E-mail: quake@stanford.edu

Published 12 January 2007, *Science* **315**, 233 (2007)
DOI: 10.1126/science.1131007

This PDF file includes:

Materials and Methods
Figs. S1 to S12
Table S1
References

Other Supporting Online Material for this manuscript includes the following:
available at www.sciencemag.org/cgi/content/full/315/5809/233/DC1

Data sets 1 to 3 (as zipped file)

Supporting Online Material

Materials and methods

Device fabrication

The device was designed in AutoCAD2004 (Autodesk, Inc.) and each layer reproduced as a chrome mask at 20,000 dpi (Fineline-Imaging). Flow molds were fabricated on 3" silicon wafers (Silicon Quest International) coated with hexamethyldisilazane (HMDS) in a vapour bath for 2 min. The wafers were then spin coated with SPR 220-7 (Shipley) initially at 500 rpm for 5 s followed by 4000 rpm for 60 s yielding a substrate height of around 6-7 μm. The molds were baked at 105 °C for 90 s followed by a 15 s I-line exposure on a MA6 contact mask aligner (Karl Suss). Next the molds were developed with 1:5 2401 developer (Microposit) in dH_2O. Finally the molds were annealed at 120°C for 20 min. Control molds were fabricated on 3" silicon wafers by spin coating SU-8 2025 (MicroChem) at 2700 rpm for 80 s followed by a 65°C bake for 2 min, 95°C for 5 min and a final step of 65°C for 2 min. The wafers were then exposed for 10 s on the I-line, followed by a post-exposure bake series of 65°C for 2 min, 95°C for 12 min and 65°C for 2 min. The wafers were then developed in SU-8 developer for 90 s followed by an acetone and isopropanol wash. One wafer from each control and flow wafer set was selected and used for all subsequent microfluidic device fabrication. The microfluidic devices were fabricated essentially as described previously (S*1*).

Linear template synthesis

Linear expression templates were generated by a two step PCR method (Fig. S2) in which the first step amplifies the target sequence and the second step adds required 5'UTR and 3'UTR for efficient ITT. Pho4 N or C-His tagged and Cbf1 N or C-His tagged versions were amplified

from yeast genomic DNA as follows: The first step PCR reaction contained 1 μM of each gene specific primer, 10 ng μL^{-1} yeast genomic DNA (SeeGene), 200 μM of each dNTP and 2.5 units of TAQ enzyme mixture (Expand High Fidelity PCR system, Roche) in a final volume of 50 μL. The reaction was cycled for 4 min at 94°C, followed by 30 cycles of 30 s at 94°C, 60 s at 53°C and 90 s at 72°C followed by a final extension of 7 min at 72°C. The products where then purified on spin columns (QIAquickPCR, Qiagen) and eluted in 75 μL of 10 mM TrisCl, pH 8.5. The purified product then served as template in the second PCR reaction using 2 μL first PCR product, 5 nM 5'ext1 primer, 5 nM 3'ext2 primer, 200 μM of each dNTP and 2.5 units of TAQ enzyme mixture (Expand High Fidelity PCR system, Roche) in a final volume of 100 μL. The reaction was cycled for 4 min at 94°C followed by 10 cycles of 30 s at 94°C, 60 s at 53°C and 90 s at 72°C followed by a final extension of 72°C for 7 min. After this first round of extension 2 μL of 5 μM 5'finalCy5 and 5 μM 3'final in dH_2O were added to each reaction and cycling was continued immediately at 94°C for 4 min followed by 30 cycles of 30 sec at 94°C, 60 sec at 50°C and 90 s at 72°C followed by a final extension of 72°C for 7 min. The final product was then purified on spin columns and eluted in 100 μL 10 mM TrisCl, pH8.5 or used directly in ITT reactions. Linear expression templates for MAX iso A, MAX iso B were synthesized essentially as above except that bacterial cDNA clones (MGC) lysed in 2.5 μL Lyse n' Go buffer (Pierce) at 95°C for 7 min where used as template in an Expand High Fidelity PCR reaction (Roche). The first PCR product was purified using the Qiaquick 96 PCR purifcation kit (Qiagen) and eluted in 80 μL of 10 mM TrisCl, pH 8.5.

To assess the fidelity of these multi-step PCR reactions and to ascertain that no point mutations accumulated during the reaction we submitted final products of MAX iso B notag, MAX iso B C-His, PHO4 C-His and CBF1 N-His to sequencing (Biotech Core). The resulting sequences showed extremely high-fidelity with no accumulation of point mutations (data not

shown).

Target DNA synthesis

All small dsDNA oligos serving as targets for transcription factor binding were synthesized by isothermal primer extension in a reaction containing 6.7 μM 5'CompCy5, 10 μM library primer, 1 mM of each dNTP, 5 units Klenow fragment (3'$\rightarrow$5' exo$^-$), 1 mM dithiothreitol 50 mM NaCl, 10 mM $MgCl_2$ and 10 mM Tris-HCl, pH7.9 in a final volume of 30 μL. All reactions were incubated at 37°C for 1 h followed by 20min at 72°C and a final annealing gradient down to 30°C at a rate of 0.1°C sec^{-1}. We added 70 μL of a 0.5% BSA dH_2O solution to each reaction and transferred the entire volume to a 384 well plate in which a 6 fold dilution series was established with final concentrations of 5'CompCy5 of 2 μM, 600 nM, 180 nM, 54 nM, 16 nM and 5 nM.

DNA arraying and device alignment

All target sequences were spotted with an OmniGrid Micro (GeneMachines) microarrayer using a CMP3B pin (TeleChem International, Inc.) for delivery onto epoxy coated glass substrates (CEL Associates). Each sample solution contained 1% BSA in dH_2O to prevent covalent linkage of the target DNA to the epoxy functional groups as well as for visualization during alignment. After spotting the arrays were quality controlled on a GenePix4000b (Molecular Devices). The arrays could then be stored in the dark at room temperature until aligned to a microfluidic device. Device alignment was done by hand on a SMZ1500 (Nikon) stereoscope and bonded overnight in the dark on a heated plate at 40°C.

Surface chemistry, protein synthesis and MITOMI

All devices were driven between 12 and 18 psi in the control line and 6 psi for the flow line. For the initial surface derivatization steps the chamber valves remained closed to prevent liquid from entering the chambers containing the spotted DNA targets (Fig. S4C). First, all accessible surface area was derivatized by flowing a solution of biotinylated BSA (Pierce) resuspended to 2 mg/mL in dH_2O for 30 min through all channels, followed by a 10 min PBS wash (Fig. S4D). Next a 500 μg/mL Neutravidin (Pierce) solution in PBS was flown for 20 min, followed by a 10 min PBS wash (Fig. S4E). Next, the "button" membrane was closed and the PBS wash continued for an additional 5 min. Then all remaining accessible surface area was passivated with the same biotinylated solution as above for 30 min, followed by a 10 min PBS wash (Fig. S4F). Finally a 1:1 solution of biotinylated-penta-histidine antibody (Qiagen) in 2% BSA in PBS was loaded for 2-5 min, after which the "button" membrane was opened and flow continued for 20 min, again followed by a 10 min PBS completing the surface derivatization procedure (Fig. S4G).

Next a standard 25 μL TNT T7 coupled wheat germ extract mixture (Promega) was prepared and spiked with 1 μL $tRNA_{Lys-bodipy-fl}$ (Promega) and 2 μL of linear expression ready template coding for the appropriate transcription factor. The mixture was immediately loaded onto the device and flushed for 5 min, after which the chamber valves were opened allowing for dead end filling of the chambers with wheat germ extract (Fig. S4H). The chamber valves were again closed and flushing continued for an additional 5 min. Next the segregation valves separating each unit cell were closed followed by opening of the chamber valves allowing for equilibration of the unit cell by diffusion. The entire device was heated to 30°C on a temperature controlled microscope stage and incubated for up to 90 min (Fig. S4I). After the incubation period the device was imaged on a modified arrayWoRx^e (AppliedPrecision) microarray scanner. Next

we performed MITOMI by closing the ”button” membrane (Fig. S4J) as well as the chamber valves (Fig. S4K) followed by a 5 min PBS wash (Fig. S4L) after which the device was imaged once more to detect the trapped molecules.

MITOMI characterization

To measure the effect of button closing rate on DNA trapping we measured the button closing rates of the 640 chamber and 2400 chamber devices. The buttons were closed at various pressures ranging from 12 psi to 24 psi in 3 psi steps to modulate button closing velocities. Movies were taken of the button closing at these various pressures using a digital camera (DFW-V500, Sony) at 25 fps. The radial button closing velocities were extracted from these videos for both devices at all pressures (Fig. S6A). Closing velocities differ between the two devices due to slight differences in architecture of the button, where the 640 chamber device had a narrower channel connecting the button with the feeding channel. In order to measure the effect of closing velocity on DNA trapping efficiency we measured the resulting ratio of trapped DNA to protein under the button after closing at various velocities (Fig. S6B). We performed all measurements at closing pressures of 15psi - 18psi reaching velocities of 4.6 μm/sec and above and thus are in a region were the closing velocity is sufficiently fast and no DNA loss is observed.

In order to assess the effectiveness of the mechanical trapping of DNA by the PDMS membrane we measured how much DNA is lost while the button is in a closed state. All experiments were performed on a 640 chamber version of the original device with a TNNNGTG library and MAX iso A C-His as the transcription factor. Two experiments were performed with various measurement intervals. The first experiment consisted of measurement intervals of 30 min for two hours followed by a final long term measurement 15 hours into the experiment. On a second device the measurement interval was extended to 60 min for four hours followed by a final

measurement at 21 hours. We then fit exponential functions to all four time-series and plotted the resulting rate constants (Fig. S7A). In order to separate the contribution of bleaching to the actual mass loss rate we plotted the measured rate constants as a function of the measurement interval and fit a linear regression, of which the intercept represents the actual mass loss rate with a value of 0.0009 sec^{-1} (Fig. S7B). We therefore observe a small mass loss of DNA from beneath the button on very long time scales, most likely due to lateral diffusion, but the loss is negligible over the time course of a normal experiment.

To ascertain the reproducibility of MITOMI we compared experiments from different days and devices for all four TFs studied (Fig. S8). All comparisons show good correlation of values including the low affinity regime. The fact that low affinity measurements correlate indicates that they do not lie near the detection limit, determined to be around 18 μM (data not shown). To arrive at a global measurement error that includes both biological as well as technical noise we compared all N and C-terminally tagged TF datasets, yielding a global measurement error of 19% (Fig. S9). Additional negative controls such as a no protein as well as a no-epitope tag MAX A control showed no non-specific trapping of DNA (data not shown).

Image and Data Analysis

All images were analyzed with GenePix3.0 (Molecular Devices). For each experiment two images where analyzed. The first image taken after the 60-90 min incubation period, was used to determine the concentration of solution phase or total target DNA concentration (Cy5 channel). The second image taken after MITOMI and the final PBS wash was used to determine the concentration of surface bound protein (FITC channel) as well as surface bound target DNA (Cy5 channel). Dissociation equilibrium constants were determined for each experiment using Prism 4 (Graphpad Software) by performing global nonlinear regression fits using a one

site binding model to the data plotted as surface bound target DNA (RFU) divided by surface protein concentration (RFU) (or effectively fractional occupancy) as a function of total target DNA concentration (RFU). The B_{max} parameter was set equal to the plateau of the consensus sequence and used for all linefits (Fig. S5). These relative K_ds (RFU^{-1}) were then transformed into absolute K_ds (M^{-1}) using a calibration curve previously established by measuring known quantities of 5'CompCy5 primer (data not shown). $\Delta\Delta$Gs were calculated with $\Delta\Delta G = RT * ln(K_d/K_{d,ref})$ at a temperature of 298 K. The highest affinity sequence was always chosen as the reference.

We estimated our measurement error by plotting affinities measured of all N-terminally tagged transcription factors versus their respective C-terminal variants. We adjusted all slopes of the linear regression fits to be uniform for all transcription factors. Our observed variance was heteroscedastic. We therefore applied a ln transform to our data which resulted in constant variance from which we obtained σ values of 0.17 and 0.40 for one and two σ respectively. Re-transforming these values yields $\hat{\sigma} = .19x$ and $2\hat{\sigma} = 0.49x$ or 19% and 49% respectively.

In silico model

Genomic binding was predicted using scripts written in Python. The yeast feature file (SGD_feature.tab downloaded on 14-11-2005) from the Saccharomyces Genome Database (SGD) was used to extract all yeast ORFs (genes) annotated as being a CDS (exon) feature from chromosomes 1-16. This filter yielded 5814 yeast genes. Regulatory sequences spanning either 500 bps (for Pho4p) or 800 bps (for Cbf1p) upstream of the start codon were extracted for each of the 5814 ORFs from yeast genomic DNA. We found that for Pho4p a 500 bps regulatory sequence is sufficient since all binding sites cluster in regions spanning the nucleosome free region between -100 to -200 bps (S2) and a second cluster centered on ≈350 bps. For Cbf1p it

was necessary to extend the regulatory sequence to -800 bps since the promoter regions were not as defined. This is likely due to the fact that Cbf1p functions as a transcriptional regulator through chromatin remodeling in the target regulatory sequence (S*3*) as opposed to binding to predefined nucloesome free regions clustered in certain regions of the regulatory sequence.

Using the determined $\Delta\Delta$G values for the flanking bases and E-box consensus halfsite or $\Delta\Delta G_{f,N}$ and $\Delta\Delta G_{c,N}$ and assuming symmetrical binding we use the same values for binding to the palindromic complements or $\Delta\Delta G_{f,X}$ and $\Delta\Delta G_{c,X}$ we can calculate full site free energy changes by adding the free energy change of each sub-site: $\Delta\Delta G_{full} = \Delta\Delta G_{f,N} + \Delta\Delta G_{c,N} + \Delta\Delta G_{c,X} + \Delta\Delta G_{c,X}$.

The following calculations were adopted from (S*4*) expressing equilibrium constants as free energies ($K_d = e^{-\Delta G/RT}$). A probability of binding (P_i) for every possible binding site i in the regulatory sequence was calculated:

$$P_i = \frac{[X]}{e^{(-\Delta G_{ref} + \Delta\Delta G_{full,i})/RT} + [X]} \tag{1}$$

where ΔG_{ref} is the free energy of transcription factor binding to its consensus sequence and $\Delta\Delta G_{full,i}$ is the change in free energy of binding to site i. [X] is the cellular concentration of transcription factor and is generally not known and thus was set equal to ΔG_{ref} yielding a P_{ref} of 0.5 for the consensus sequence. For Cbf1p the cellular concentration is estimated to be 6890 molecules (S*5*) or assuming a cell size of 2 pL approximately 5.5 nM, which corresponds well with the measured K_d of 16.63 nM. Since [X] was set equal to ΔG_{ref}, P_i could be expressed solely as a function of the change in free energy $\Delta\Delta G_{full,i}$:

$$P_i = \frac{1}{e^{\Delta\Delta G_{full,i}/RT} + 1} \tag{2}$$

Now a probability of occupancy (P_{occ}) was calculated for each ORF regulatory sequence:

$$P_{occ} = 1 - \prod_{i=1}^{windows} \left(\frac{1}{1 + e^{(-\Delta\Delta G_{full,i})/RT}} \right) \tag{3}$$

We then used a P_{occ} cutoff of 0.2 for both Cbf1p and Pho4p selecting genes with regulatory sequences likely to be bound by their respective transcription factors. This resulted in 55 genes predicted to be regulated by Pho4p and 35 genes regulated by Cbf1p.

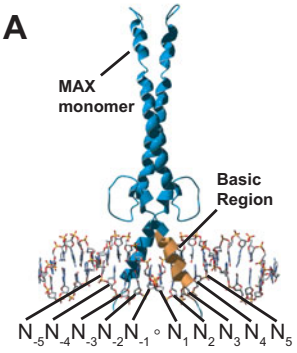
To reduce the false positive rate and understand how well the binding sites are conserved across species, we recalculated the P_{occ}s for all selected genes in yeast sensu stricto species, including the Washington University sequences for *S. bayanus, S. castellii, S. kluyveri, S. kudriavzevii, S. mikatae* (S*6*) and the MIT sequences for *S. mikatae, S. bayanus, S. paradoxus* (S*7*). From these values we calculated a conservation score by dividing the number of species that returned a $P_{occ} > 0.2$ by the total number of species that returned a value above zero. For species *S. mikatae* and *S. bayanus* for which two sequences were available we collapsed them using a logical disjunction (OR). We then used a cutoff of 0.25, or a 25% conservation across species, to select the final gene sets. This process did not substantially change the gene predictions, although the sets were slightly smaller, as described in the main text.

MIPS Funcat

All functional annotations and enrichments were determined using the MIPS FunCat server (S*10*). The significance cutoff was left at a p-value of $5e{-3}$.

Supporting Figures

A



B

	Basic Region												
bHLH TF	1	2	3	4	5	6	7	8	9	10	11	12	13
Pho4p	K	R	E	S	H	K	H	A	E	Q	A	R	R
Cbf1p	R	K	D	S	H	K	E	V	E	R	R	R	R
MAX	K	R	A	H	H	N	A	L	E	R	K	R	R
Consensus					H				E			R	R

C

	5'				E-Box								3'	
Name	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7	Members
NNCAC	T	T	N	N	C	A	C	G	T	G	T	T	T	T	16
NNNCAC	T	N	N	N	C	A	C	G	T	G	T	T	T	T	64
NNNNGTG	T	T	G	N	N	N	N	G	T	G	G	G	T	G	256
CACNNN	T	T	G	T	C	A	C	N	N	N	A	C	T	T	64
GTGNNN	T	T	T	T	C	A	C	G	T	G	N	N	N	T	64
													Sum		464

Fig. S1. (A) Atomic structure of a MAX homodimer bound to a target DNA sequence (S*12*). The basic region of one monomer is colored in gold. The bases are labeled according to the nomenclature used in the text. (B) Alignment of the basic regions of the transcription factors Pho4p, Cbf1p and MAX. (C) Summary of the various target DNA libraries synthesized.

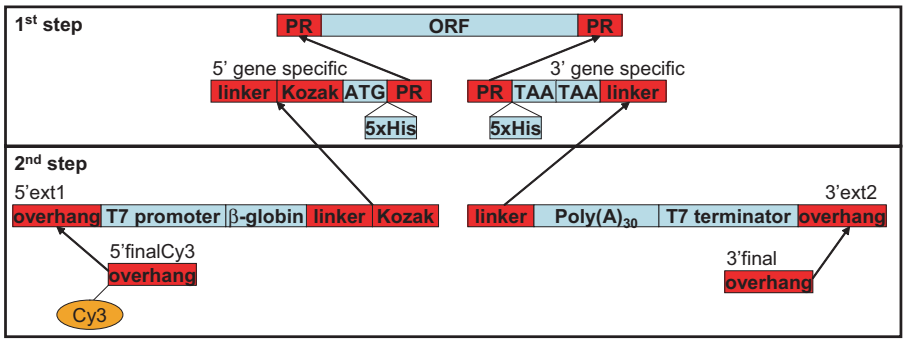


Fig. S2. Overview of the 2 step PCR method for generating linear expression ready templates to be used in ITT. An ORF obtained from either yeast genomic DNA or a cDNA clone serves as the template being amplified in the first step with gene specific primers. These primers may carry sequences coding for an epitope tag of choice, here either a N or C terminal penta-histidine. In the second PCR step product of the first step is used as template, extended with extension primers carrying all required 5' and 3'UTR sequences for efficient ITT expression. The full length product is then amplified using final primers carrying moieties such as fluorophores or biotin.

Figure S3

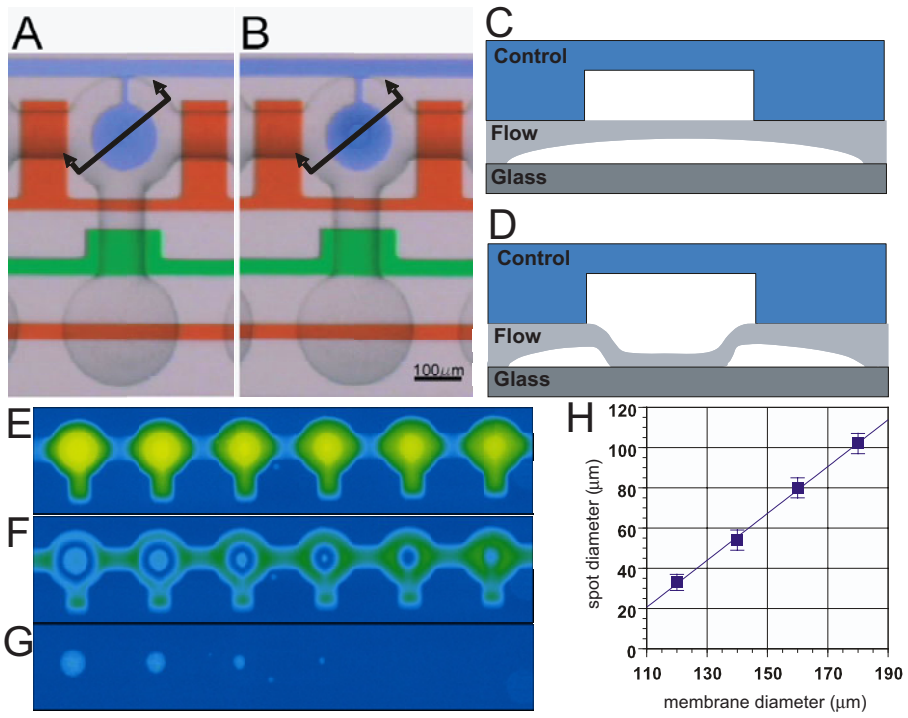


Fig. S3. Overview of the button membrane. (A-B) optical micrographs of one unit cell (see Fig. 1) with a button in the open (A) and closed state (B). (D-C) Cross-sectional schematics of the regions indicated in (A-B). (E-H) A device was designed with various button diameters to assess how button diameter effects the resulting spot diameter. (E-F) Fluorescence intensity maps showing target DNA concentration. The button diameter decreases from the left to right. The two rightmost buttons failed to contact the surface. (H) Analysis of the data showing how button diameter influences spot diameter. The resulting spot diameters have very low variability and the smallest spot generated in this experiment was ≈ 30 μm.

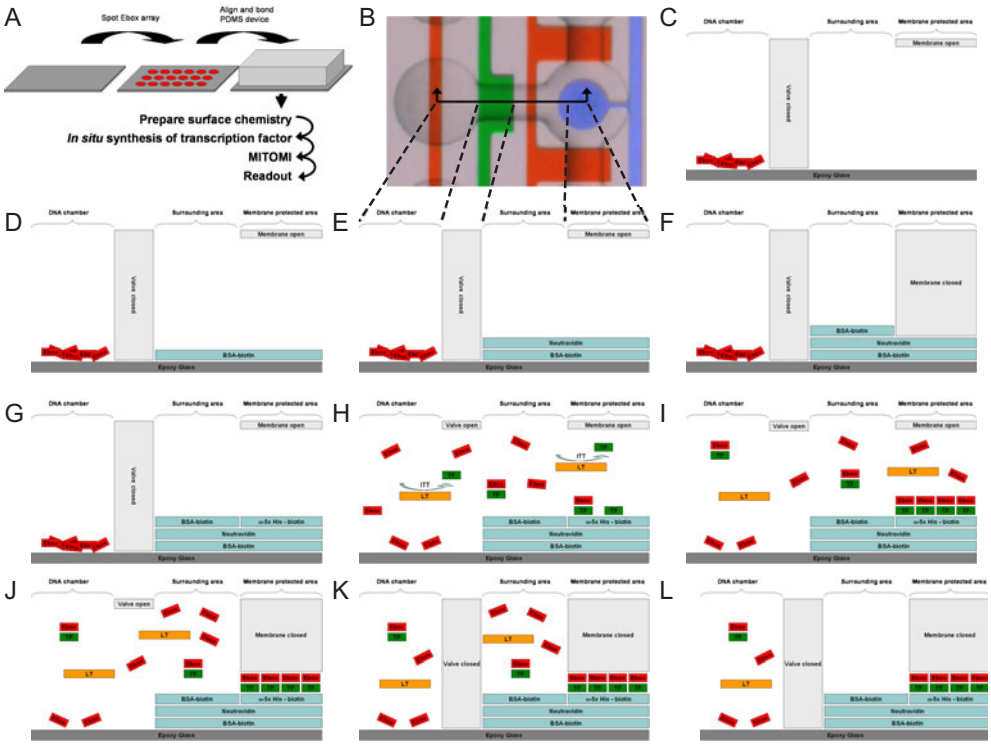


Fig. S4. (A) Overview of the experimental approach starting with a plain epoxy substrate to be spotted with 2400 spots of a target DNA library. The finished microarray is then aligned and bonded to one of our microfluidic devices after which the surface is prepared, protein synthesized and MITOMI performed. (B) Micrograph of one of the microfluidic unit cells, shown here again for reference. The dashed lines show which regions of the unit cell are schematically depicted in panels C-L. (C) Before any fluid is introduced into the device the chamber valve (green channel in panel B) is closed to prevent flooding of the DNA chamber. (D) Next biotinylated BSA is introduced into our device which covalently bonds to the epoxy functional groups, both activating (via the biotin moieties) and passivating (epoxy groups) the surface. (E) A solution of neutravidin is introduced forming a monolayer on top of the biotinylated BSA layer. (F) The ”button” membrane is closed to protect the detection area from passivation via biotinylated BSA which passivates all accessible surface area. (G) Any unbound biotinylated BSA is purged before the ”button” membrane is opened again allowing access to the neutravidin surface below to which a biotinylated penta-histidine antibody is attached, concluding the surface derivatization. (H) ITT programmed with linear expression template is introduced into the device and allowed to flood the DNA chamber causing the solvation of the stored target DNA. Transcription factor is being synthesized and is pulled down to the surface by penta-histidine antibody. (I) The synthesized transcription factors functionally interact with the solvated target DNA pulling it down to the surface as well. (J) After 60-90 min the ”button” membrane is closed again mechanically trapping any molecular interactions taking place on the surface allowing all solution phase molecules to be washed away without loss of surface bound material (K-L).

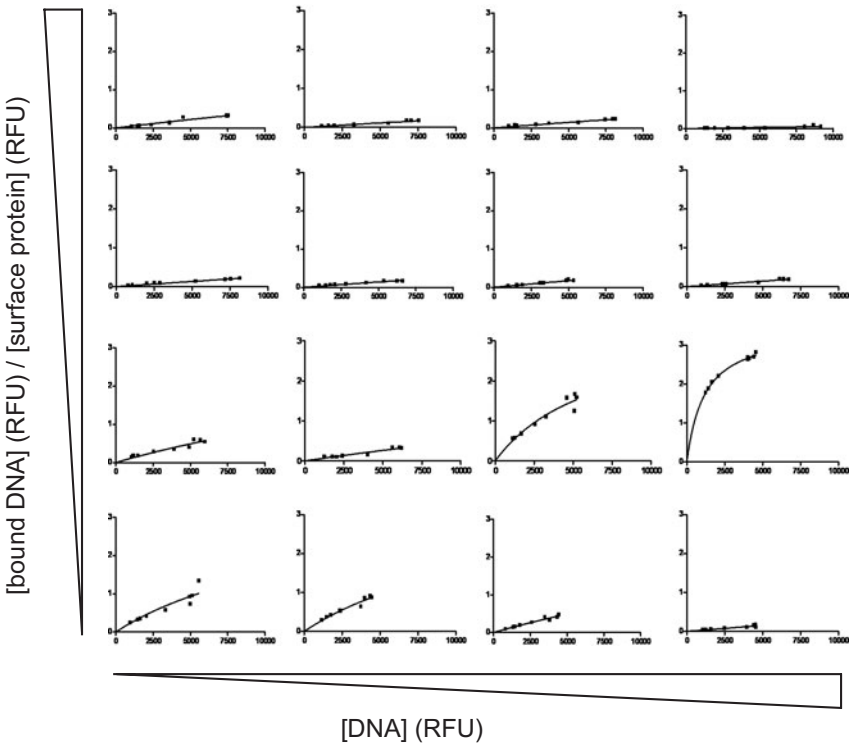


Fig. S5. Raw data from which K_ds are determined. 16 representative binding curves from the MAX isoform A C-his dataset.

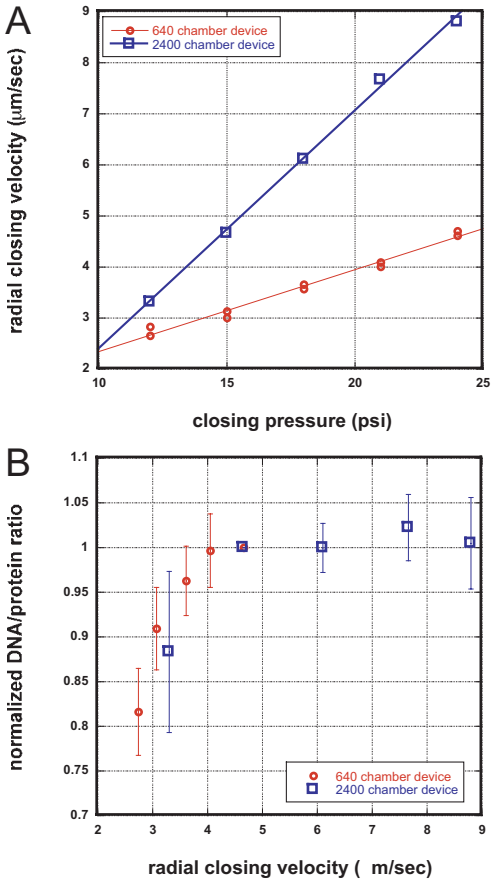


Fig. S6. Effect of button closing velocity on trapping efficiency. (A) Radial closing velocities were measured at various pressures on two different devices, with slightly different architectures. (B) The effect on the above closing velocities on trapping efficiencies. Below 4 μm/sec trapping is strongly dependent on button closing velocity. This dependence disappears above 4 μm/sec where the response plateaus off.

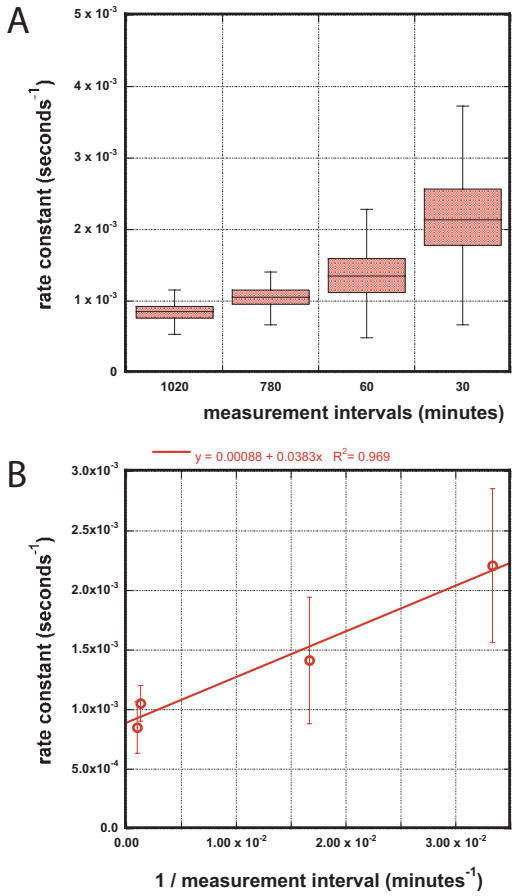


Fig. S7. Measurements illustrating the effectiveness of the mechanical trapping of molecules. (A) Rate constants determined at four measurement intervals. (B) The same rate constant averages as in (A) plotted as a function of the measurement interval. The dependence on the measurement interval indicates that bleaching contributes to the measured rate constants. The intercept represents the actual mass loss rate (without the contribution of bleaching) with an average value of 0.0009 s^{-1}.

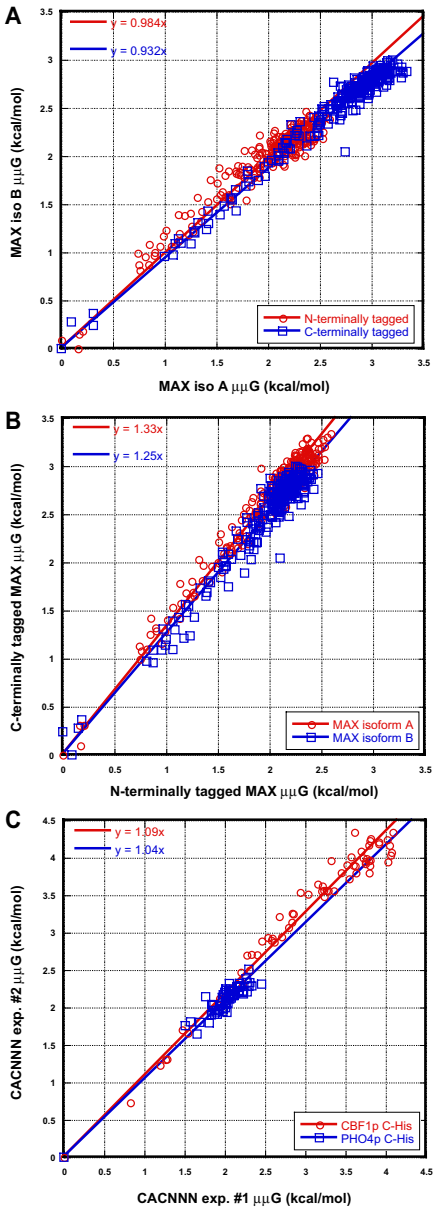


Fig. S8. MAX dataset comparisons. (A) Comparison of the two MAX isoforms. The slope of the linear regression fit is close to one for both the N and C-terminally tagged versions and the values remain correlated throughout indicating that there is no difference between the isoforms. (B) Comparison of the two epitope tag locations. Here an increase in the slope is visible indicating that C-terminally tagged versions exhibit increased affinity to their target sequences. (C) Comparison of the day to day variability. Datasets taken for Pho4p and Cbf1p taken on different days and on different devices are compared to one another.

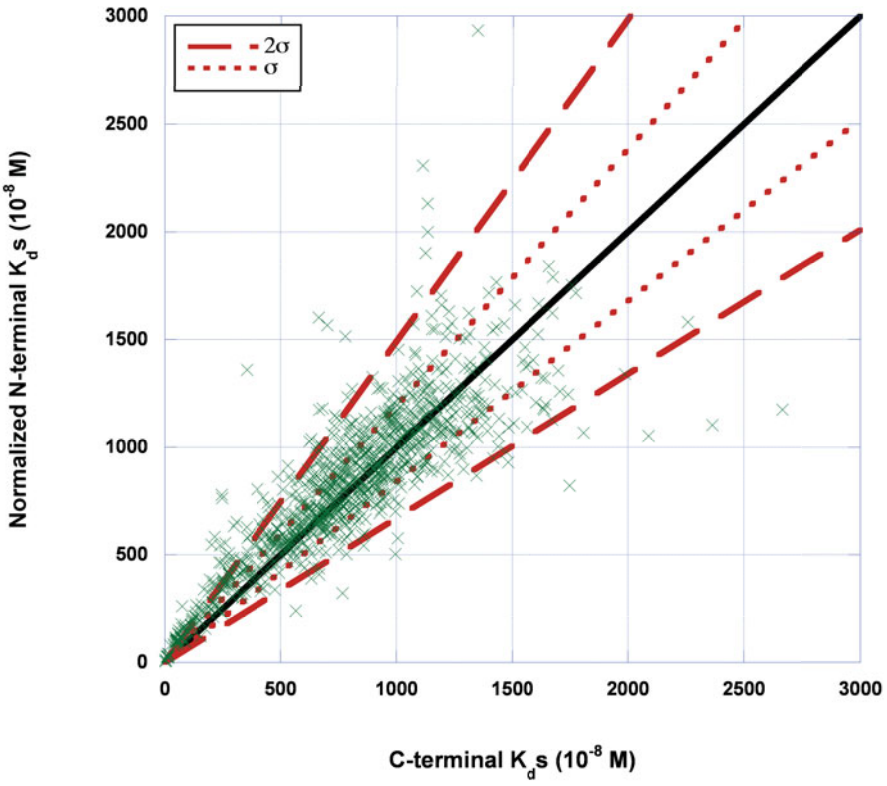


Fig. S9. Scatter plot of all binding affinities to sequences $N_{-4} - N_{-1}$ determined for all transcription factors N and C terminally tagged. The four N-terminal datasets were adjusted to have the same slope as the C-terminal data. The dotted and dashed red lines show one and two sigma of the spread in the data at 19% and 49% respectively.

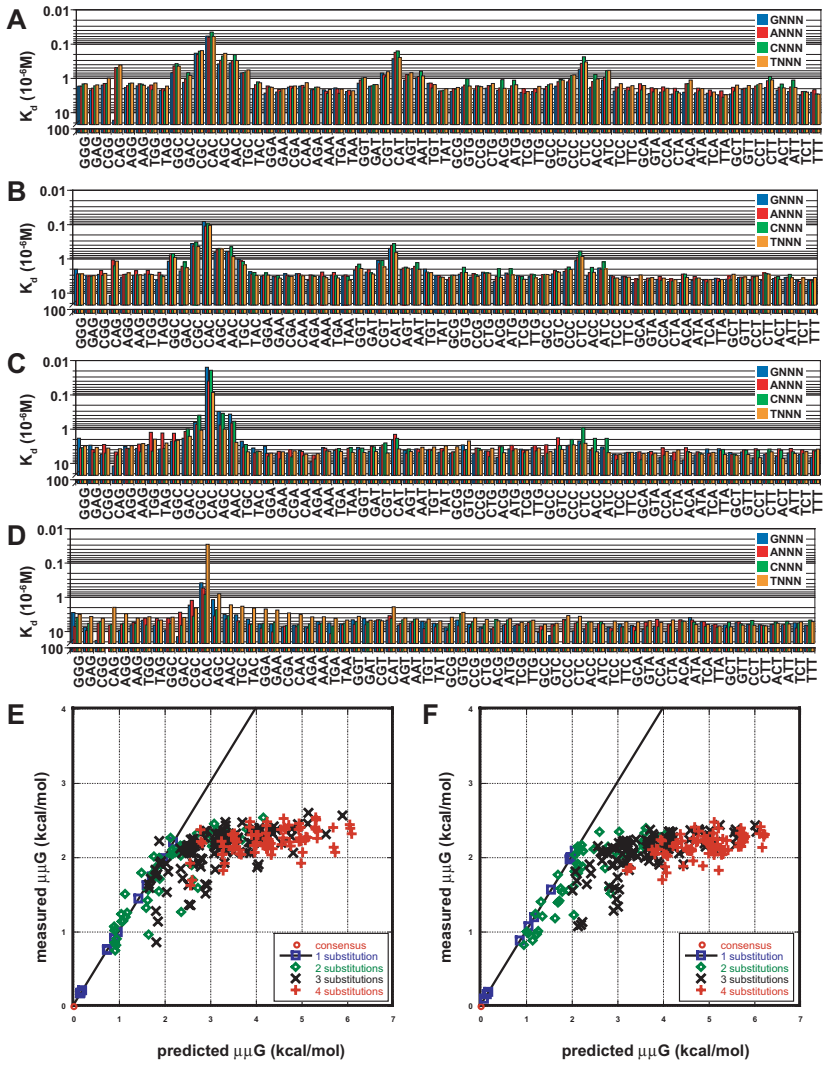


Fig. S10. Same as Fig. 2 except that all transcription factors are N-terminally his tagged.

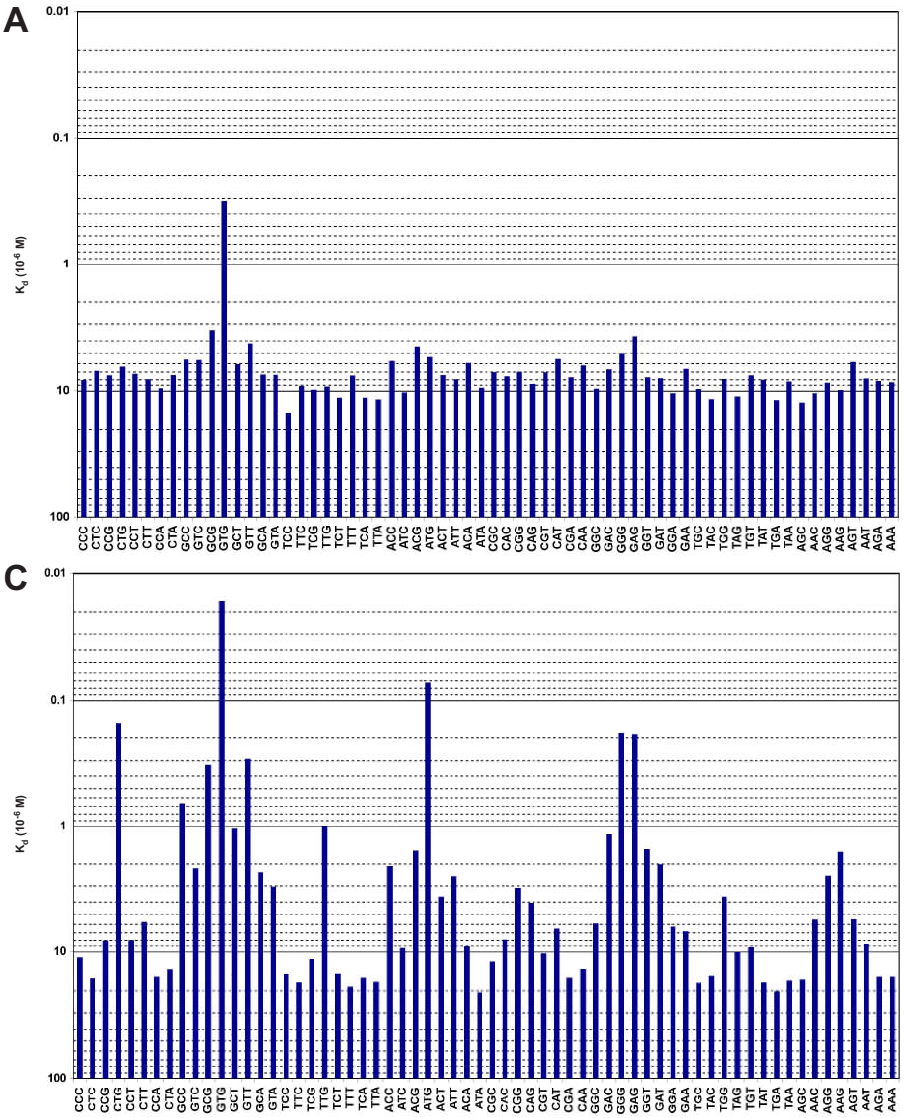


Fig. S11. Data of binding affinities for sequences covering the 3' half-site $N_1 - N_3$ for Pho4p (A) and Cbf1p (B).

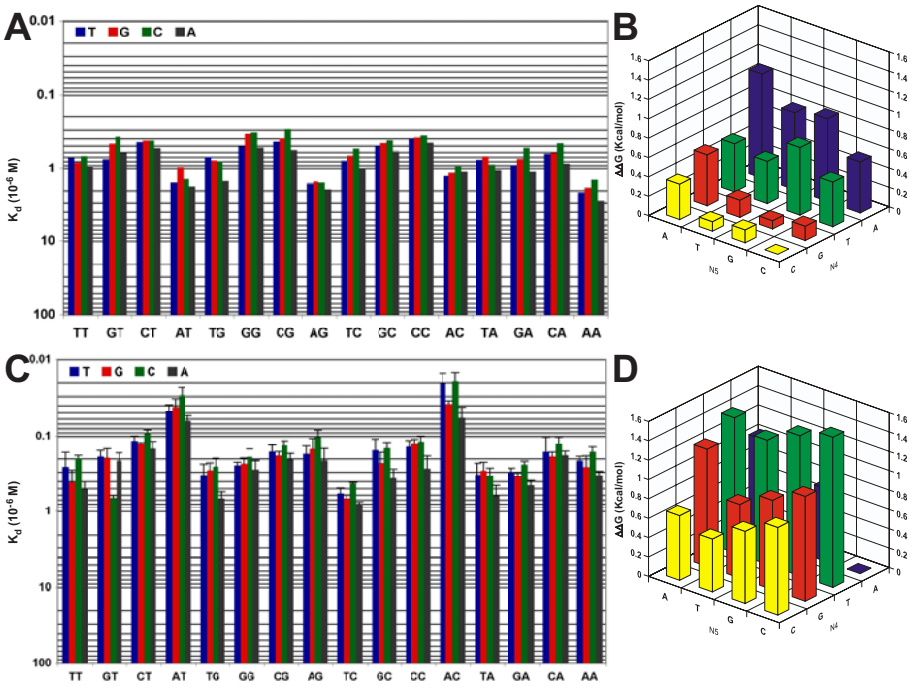


Fig. S12. Binding affinities of Pho4p (A) and Cbf1p (C) for the 3' flanking bases $N_4 - N_6$. (B,D) affinities of Pho4p and Cbf1p to flanking bases ($N_4 - N_5$). Error bars in (C) show one standard deviation and were included to show the significance of the difference in binding affinity of the 3rd flanking base N_6.

Supporting Tables

Pho4p

FUNCTIONAL CATEGORY	P-VALUE
34.01.03 homeostasis of anions	1.25E-08
34.01.03.03 homeostasis of phosphate	1.06E-07
20.01.01.07 anion transport (Cl, SO4, PO4, etc.)	2.63E-07
20.01.01.07.07 phosphate transport	2.76E-07
01.04 phosphate metabolism	1.72E-06
01.04.01 phosphate utilization	4.91E-05
42.25 vacuole or lysosome	9.95E-05
01.05.01 C-compound and carbohydrate utilization	2.33E-04
34.01 ionic homeostasis	4.82E-04
01 METABOLISM	6.35E-04
34 INTERACTION WITH THE CELLULAR ENVIRONMENT	8.56E-04
20.01 transported compounds (substrates)	1.28E-03
20.01.01 ion transport	1.34E-03
01.05 C-compound and carbohydrate metabolism	1.63E-03
14.07.02.01 O-directed glycosylation	3.11E-03
20 CELLULAR TRANSPORT, TRANSPORT FACILITATION AND TRANSPORT ROUTES	3.34E-03

Cbf1p

FUNCTIONAL CATEGORY	P-VALUE
40 CELL FATE	3.25E-05
40.01 cell growth / morphogenesis	1.47E-04
42 BIOGENESIS OF CELLULAR COMPONENTS	4.15E-04
43.01.03.05 budding, cell polarity and filament formation	6.56E-04
10.03.01.01.09 G2/M transition of mitotic cell cycle	6.98E-04
43 CELL TYPE DIFFERENTIATION	7.62E-04
43.01 fungal/microorganismic cell type differentiation	7.62E-04
43.01.03 fungal and other eukaryotic cell type differentiation	7.62E-04
14.07 protein modification	9.14E-04
10.03.01.01 mitotic cell cycle	2.51E-03
01.02.01.14 conjunction of sulfate	3.56E-03
14.07.09 posttranslational modification of amino acids (e.g. hydroxylation, methylation)	3.75E-03
14.07.03 modification by phosphorylation, dephosphorylation, autophosphorylation	3.88E-03
14 PROTEIN FATE (folding, modification, destination)	4.01E-03

Table S1. Functional enrichments of our Pho4p and Cbf1p gene sets determined using the MIPS-FunCat server. Blue entries indicate functions related to phosphate or amino acid metabolism for Pho4p and Cbf1p, respectively. Red entries are functions involved in cell cycle control and budding.

References

S1. T. Thorsen, S. J. Maerkl, S. R. Quake, *Science* **298**, 580 (2002).

S2. G. C. Yuan, *et al.*, *Science* **309**, 626 (2005).

S3. N. A. Kent, S. M. Eibert, J. Mellor, *J Biol Chem* **279**, 27116 (2004).

S4. J. A. Granek, N. D. Clarke, *Genome Biol* **6**, R87 (2005).

S5. S. Ghaemmaghami, *et al.*, *Nature* **425**, 737 (2003).

S6. P. Cliften, *et al.*, *Science* **301**, 71 (2003).

S7. M. Kellis, N. Patterson, M. Endrizzi, B. Birren, E. S. Lander, *Nature* **423**, 241 (2003).

S8. G. Thijs, *et al.*, *J Comput Biol* **9**, 447 (2002).

S9. J. D. Hughes, P. W. Estep, S. Tavazoie, G. M. Church, *J Mol Biol* **296**, 1205 (2000).

S10. A. Ruepp, *et al.*, *Nucleic Acids Res* **32**, 5539 (2004).

S11. N. Ogawa, J. DeRisi, P. O. Brown, *Mol Biol Cell* **11**, 4309 (2000).

S12. A. R. Ferre-D'Amare, G. C. Prendergast, E. B. Ziff, S. K. Burley, *Nature* **363**, 38 (1993).